

AP5/B1015
3/30/92

d his

(FILE 'USPAT' ENTERED AT 14:52:45 ON 30 MAR 92)

L1 161 S PDGF OR PLATELET DERIVED GROWTH FACTOR
L2 2545 S ENDOTHEL? OR FIBROBLAST
L3 106 S L1 AND L2
L4 1 S L3 AND 349
L5 2 S (36 OR 37 OR 38)(2W)(KD OR KILODALTON?) AND L3
L6 57 S (36 OR 37 OR 38)(2W)(KD OR KILODALTON?)
L7 67 S (36,000 OR 37,000 OR 38,000)(W)(MR OR MOLECULAR(W)WEIGHT
OR
L8 60 S L7 NOT L6
L9 117 S L6 OR L8
L10 3 S L9 AND L1

=> d ti ab in pddfd ccls 1-3

US PAT NO: 5,051,364 [IMAGE AVAILABLE] L10: 1 of 3
TITLE: Anti-lipocortin-I and anti-lipocortin-II monoclonal
antibodies

ABSTRACT:

The present invention provides monoclonal antibodies which are specific for one but not both of human lipocortin-I and human lipocortin-II, as well as cultures of hybridomas and other types of cells producing such antibodies.

INVENTOR: Clare M. Isacke, La Jolla, CA
Ian S. Trowbridge, San Diego, CA
Tony Hunter, San Diego, CA

DATE ISSUED: Sep. 24, 1991

DATE FILED: Dec. 20, 1989

US-CL-CURRENT: 435/240.27; 424/85.8, 85.91; 435/70.21; 530/387, 389

US PAT NO: 4,935,233 [IMAGE AVAILABLE] L10: 2 of 3
TITLE: Covalently linked polypeptide cell modulators

ABSTRACT:

Described is a new class of polypeptide cell modulators characterized by being composed of two covalently linked cell modulators in a linear polypeptide sequence. Such dual function polypeptides have new and particularly useful activities when the component polypeptide cell modulators are interferons, lymphokines or cytotoxins which act through different and specific cell receptors to initiate complementary biological activities.

INVENTOR: Leslie D. Bell, Thame, United Kingdom
Keith G. McCullagh, Princes Risborough, United Kingdom
Alan G. Porter, High Wycombe, United Kingdom

DATE ISSUED: Jun. 19, 1990

DATE FILED: Dec. 2, 1985

US-CL-CURRENT: 424/85.5, 85.6, 85.7; 435/69.51; 530/351

US PAT NO: 4,543,439 L10: 3 of 3
TITLE: Production and use of monoclonal antibodies to
phosphotyrosine-containing proteins

ABSTRACT:

A hybridoma cell line is disclosed that secretes monoclonal antibodies

which serve as a high titer, reproducible, biological reagent useful in biological/medical research for isolating and identifying phosphotyrosine-containing proteins. In addition, the antibodies have potential uses in diagnosis of a variety of diseases, including certain cancers. The antibodies, which have demonstrated affinity for a variety of molecules containing o-phosphotyrosine residues, were prepared using a synthetic analog, p-azobenzyl phosphonate (ABP) covalently linked to a carrier protein, as the antigen.

INVENTOR: A. Raymond Frackelton, Jr., East Providence, RI
Herman M. Eisen, Waban, MA
Alonzo M. Ross, Bensalem, PA

DATE ISSUED: Sep. 24, 1985

DATE FILED: Dec. 13, 1982

US-CL-CURRENT: 435/70.21, 240.27, 948; 935/92

=>

2/7/1

9828438 BIOSIS Number: 93013438

DIFFERENTIAL BINDING BIOLOGICAL AND BIOCHEMICAL ACTIONS OF RECOMBINANT PDGF AA AB AND BB MOLECULES ON CONNECTIVE TISSUE CELLS

GROTEENDORST G R; IGARASHI A; LARSON R; SOMA Y; CHARETTE M
DEP. CELL BIOLOGY ANATOMY, UNIV. MIAMI SCH. MED., P.O. BOX 016260, MIAMI, FLORIDA 33101.

J CELL PHYSIOL 149 (2). 1991. 235-243. CODEN: JCLLA

Full Journal Title: Journal of Cellular Physiology

Language: ENGLISH

We have compared the biological and biochemical properties of recombinant PDGF AA, AB, and BB using three types of fibroblastic cells: NIH/3T3, human skin fibroblast, and fetal bovine aortic smooth muscle. PDGF binding, receptor autophosphorylation, phosphatidyl inositol hydrolysis, as well as chemotactic and mitogenic responses of the cells were analyzed. PDGF-AB and PDGF-BB showed similar receptor binding, receptor autophosphorylation, and potent biological activity for all three of the cell types tested. In contrast, PDGF-AA was biologically active only for the NIH/3T3 cells in which binding sites for PDGF-AA were abundant, but was inactive for bovine aortic smooth muscle cells and human skin fibroblasts in which binding sites for PDGF-AA were absent. PDGF-AA could not induce any biochemical changes in the human skin fibroblasts or smooth muscle cells. Western blot studies with anti-Type .alpha. and .beta. PDGF receptor antibodies indicate that the NIH/3T3 cells contained both PDGF .alpha. and .beta. receptors, whereas the human skin fibroblasts and bovine smooth muscle cells contained only detectable levels of .beta. receptors. These results indicate that cells possessing high levels of PDGF .beta. receptors only are capable of responding equally well to either PDGF AB or BB.

2/7/2

6951686 BIOSIS Number: 87012287

KIDNEY EPITHELIAL CELLS EXPRESS C-SIS PROTOONCOGENE AND SECRET PDGF-LIKE PROTEIN

KARTHA S; BRADHAM D M; GROTEENDORST G R; TOBACK F G
DEP. MED., BOX 453, UNIV. CHICAGO, 5841 SOUTH MARYLAND AVE., CHICAGO, ILL. 60637.

AM J PHYSIOL 255 (4 PART 2). 1988. F600-F606. CODEN: AJPHA

Full Journal Title: American Journal of Physiology

Language: ENGLISH

Nontransformed monkey kidney cells (BSC-1 line), used as a model for renal epithelium, were assayed for release of platelet-derived growth factor (PDGF)-like proteins. BSC-1 cells continuously released a mitogenic activity for fibroblasts and a chemoattractant activity for smooth muscle cells, each of which was inhibited 80-90% by an antibody to human PDGF. A cDNA probe for the PDGF B-chain gene (c-sis), but not for the A-chain gene, hybridized to mRNA obtained from growing and quiescent cells. c-sis gene expression and PDGF-like protein secretion were studied in the presence of known growth-regulatory molecules. A secreted BSC-1 cell protein identical

transforming growth factor .beta.2 inhibited DNA synthesis in growing cultures and induced marked accumulation of c-sis mRNA without a corresponding increase in the release of PDGF-like activity. Adenosine diphosphate stimulated DNA synthesis in quiescent cultures and enhanced both c-sis expression and release of PDGF-like activity. However, growing the quiescent cells did not express the PDGF receptor gene or exhibit a mitogenic response to authentic PDGF. Thus the PDGF-like protein released by these kidney epithelial cells could contribute to growth control by a paracrine mechanism.

?s pdgf(w)like not s2

1829 PDGF
134382 LIKE
131 PDGF(W)LIKE
2 S2

S3 130 PDGF(W)LIKE NOT S2

?s s3 and (endothel? or fibroblast?)

130 S3
43997 ENDOTHEL?
65910 FIBROBLAST?

S4 61 S3 AND (ENDOTHEL? OR FIBROBLAST?)

?s s4 and (33994W)(daaoooramóóóó) or 336(2w)(kd or kilodalton?) or (36000(2n)molecular) or ((36000(2n)(molecular or relative or dalton?))

>>>Command is too complex--please simplify

?s s4 and (349 or 36 or 36000)

61 S4
853 349
66080 36
78 36000

S5 2 S4 AND (349 OR 36 OR 36000)

?t s5/6/all

5/6/1

9100187 BIOSIS Number: 93085187

EFFECTS OF T CELLS ON PLATELET-DERIVED GROWTH FACTOR-LIKE PROTEIN SECRETION FROM ENDOTHELIAL CELLS

5/6/2

7179912 BIOSIS Number: 88102657

RETINAL PIGMENT EPITHELIAL CELLS PRODUCE PDGF-LIKE PROTEINS AND SECRETE THEM INTO THEIR MEDIA

?t s5/7/all

5/7/1

9100187 BIOSIS Number: 93085187

EFFECTS OF T CELLS ON PLATELET-DERIVED GROWTH FACTOR-LIKE PROTEIN SECRETION FROM ENDOTHELIAL CELLS

SHADDY R E; HANSEN J C; COWLEY C G

100 N. MEDICAL DRIVE, SALT LAKE CITY, UTAH 84113.

J HEART LUNG TRANSPLANT 11 (1 PART 1). 1992. 48-56. CODEN: JHLTE

Language: ENGLISH

The cause of accelerated graft arteriosclerosis after heart transplantation is unknown. To examine whether the interactions of T cells and endothelial cells (ECs) could contribute to the cause of this phenomenon, T cells were co-incubated with human umbilical artery endothelial cells (HUAEC) and human vein endothelial cells (HUVEC) and the resulting conditioned medium supernatant was assayed for the presence of platelet-derived growth factor (PDGF)-like protein. PDGF-like protein secretion was significantly greater from HUAECs co-incubated with T cells at T-cell/HUAEC ratios of 30:1 (6.9 \pm 1.1 fmol/10⁶ ECs) and 10:1 (6.0 \pm 1.1 fmol/10⁶ ECs) than the combined background secretion from HUAECs and T cells incubated separately (2.6 \pm 0.5 fmol/10⁶ ECs).

Check Ref

PDGF-like protein secretion above background levels from HUAECs was significantly greater, however, than from HUVECs co-incubated with T cells at T-cell/HUVEC ratios of 30:1 (1.0 \pm 0.4 fmol/10⁶ ECs) and 10:1 (0.75 \pm 0.36 fmol/10⁶ ECs) (p < 0.05). In four experiments, preincubation of HUAECs with γ -interferon induced HLA-DR antigen expression but actually caused a decrease in T-cell-induced PDGFR-like protein secretion above background levels (3.0 \pm 0.6 fmol/10⁶ ECs) when compared to nonstimulated HUAECs (4.0 \pm 0.4 fmol/10⁶ ECs; p < 0.05). PDGF-like protein secretion was minimal at 1 hour and increased over time to a maximum at 24 hours. The conclusion is that T cells are capable of inducing secretion of a very potent mitogen, pdgf-like protein, from endothelial cells. This cell-inducing production of PDGF-like protein appears to be primarily or exclusively a property of arterial endothelial cells.

5/7/2

7179912

Biosis Number: 88102657

RETINAL PIGMENT EPITHELIAL CELLS PRODUCE PDGF-LIKE PROTEINS AND SECRETE THEM INTO THEIR MEDIA

CAMPOCHIARO P A; SUGG R; GROTENDORST G; HJELMELAND L M

DEP. OPHTHALMOL., UNIV. VA. SCH. MED., CHARLOTTESVILLE, VA. 22908.

EXP EYE RES 49 (2) 1989 217-228. CODEN: EXERA

Full Journal Title: Experimental Eye Research

Language: ENGLISH

Get!

Human retinal pigment epithelial cells at confluence was used to condition serum-free Dulbecco's modified Eagle's medium. Conditioned media were exhaustively dialyzed against 0.5 M acetic acid, lyophilized, and subjected to Western blot analysis, using as primary antibody an IgG fraction prepared from goat antiserum directed against human platelet-derived growth factor. Native platelet-derived growth factor was resolved as a band with Mr of 30 kDa under non-reducing conditions, while bands with Mr of 36-38 kDa and 18.5 kDa were resolved from retinal pigment epithelial cell-conditioned media. Acid extracts of retinal pigment epithelial cells also contained bands at 36-38 kDa and media conditioned for 48 hr exhibited much denser bands than media conditioned for 24 hr. No bands were detected when non-immune goat IgG fractions were substituted for primary antibody and when conditioned media were prepared from several human fibroblast lines in the same manner as those prepared from retinal pigment epithelial cells, no detectable bands or only a faint shadow at 36 kDa were seen. Retinal pigment epithelial cell-conditioned media prepared in the presence of [³⁵S]methionine were loaded on an anti-platelet-derived growth factor IgG affinity column, eluted, and subjected to SDS-polyacrylamide gel electrophoresis. Bands with Mr slightly less than 36 kDa and 18 kDa were visualized by autoradiography, demonstrating that the platelet-derived growth factor-like proteins in retinal pigment epithelial cell-conditioned media are newly synthesized. Two fractions eluted from the column also markedly stimulated fibroblast chemotaxis and incorporation of [³H]thymidine, both of which were neutralized by soluble platelet-derived growth factor IgG. These data suggest that retinal pigment epithelial cells in culture produce platelet-derived growth factor-like proteins and secrete them into their media where they are capable of stimulating fibroblast chemotaxis and proliferation.

30mar92 15:25:56 User217743 Session D44.3

\$10.44 0.116 Hrs File5

\$0.00 2 Type(s) in Format 6

\$3.00 4 Type(s) in Format 7

\$3.00 6 Types

\$13.44 Estimated cost File5

\$1.25 DIALNET

\$14.69 Estimated cost this search

\$22.33 Estimated total session cost 0.258 Hrs.

logoff: level 29.01.02 D 15:25:56

DIALNET: call cleared by request

```

1980      1990      2000      2010      2020      2030      2040
AGATAGAATGACAGTCCGTCAAACAGATTGTTTGCAAAGGGGAGGCAT----CAGTGT--CTTGGCAGGCT
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
AGGTACTAGCTGAGGTTATTTTAAAAG-----CAGCAGTGTGCCTACTTTTGGAGTGTAAACCGGGGAGG-G
3240      3250      3260      3270      3280      3290      3300

      2050      2060      2070      X
GATTTCTAGGTAGGAAATG-TGGTAG-CTCACG
| | | | | | | | | | | | | | | | | |
AAATTATAG-----CATGCTTGCAGACAGACCTGCTCTAGCG
3310      3320      3330      3340

```

3. US-07-752-427-1 (1-2075)

HUMFGFAA Human fibroblast growth factor receptor mRNA, comp

```

LOCUS      HUMFGFAA      3901 bp ss-mRNA      PRI      15-MAR-1991
DEFINITION Human fibroblast growth factor receptor mRNA, complete cds.
ACCESSION  M60485
KEYWORDS
SOURCE      Human, cDNA to mRNA, clone flg 5.
ORGANISM    Homo sapiens
             Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
             Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominidae.
REFERENCE   1 (bases 1 to 3901)
AUTHORS     Kiefer,M. C. , Baird,A. , George-Nascimento,C. , Nguyen,T. , Mason,O. B. ,
             Boley,L. J. , Valenzuela,P. and Barr,P. J.
TITLE       Molecular cloning of a three-immunoglobulinlike-domain form of a
             human basic fibroblast growth factor receptor cDNA: Expression of a
             biologically active extracellular domain in a baculovirus system
JOURNAL      Unpublished (1991)
STANDARD    simple staff_entry
FEATURES
  CDS        Location/Qualifiers
             672..3134
             /product="fibroblast growth factor receptor"
             /gene="flg"
             /codon_start=1
  sig_peptide 672..734
             /gene="flg"
             /codon_start=1
  mat_peptide 735..3131
             /gene="flg"
             /codon_start=1
BASE COUNT  839 a   1187 c   1139 g   736 t
ORIGIN

```

```

Initial Score      =      146   Optimized Score      =      878   Significance      =      0.00
Residue Identity   =      47%   Matches              =      1050   Mismatches       =      912
Gaps               =      226   Conservative Substitutions      =      0

```

```

      X      10      20      30      40      50
      CCCGGC-CGA-CAGCCCCGAGACGACAGCCCGGCGCGTC-CCGGTCCCCACCTCCGACCACC
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
GCGAGCGCGCCGCTGCTTGAAAAGCCGCGGAACCCAAG---GACTTTTCTCCGGT-CCGAGCTCGGGGCGCC
170      180      190      200      210      220      230

60      70      80      90      100      110      120
GC-CA-GCGCTCAGGCCCGCGCTCCCGCTCG--CCGCCACCGCGCCCTCCGCTCCGCCCGCAGTGCCAA
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
CCGCAGGCGC-ACGGTACCGTGCT-GCAGCTGGGCACGCCGCGGCGCCGGGGCCTCCGCAGGC---GCCGG
240      250      260      270      280      290      300

130      140      150      160      170      180      190
CCATGACCGCCGCCAGTATGGG●C---GTCCGCGTCGCCTTCGTGGTC●CCTCGCCCTCTGCAGCCGGC
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
CC-TGCGTTCTGGAGGAGGGGGGCACAAGGTCTGGAGACCCCGGGTGG-CGGACGGGAGCCCTCCCCCGCC

```


480 490 500 510 520 530 540
TTCCAGAGCAGCTGCAAGTACCAAGTGCACGTGCCTGGACGGGGCGGTGGGCTGCATGCCCCTGTGCAGCATG
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
TTCCAAGCAGCTGCAAATACCAATGCACTTGCCTGGATGGGGCCGTGGGCTGCGTGCCCCTATGCAGCATG
1740 1750 1760 1770 1780 1790 1800

550 560 570 580 590 600 610
GACGTTCTGTCTGCCAGCCCTGACTGCCCCTTCCCGAGGAGGGTCAAGCTGCCCGGGAAATGCTGCGAGGAG
||||| || ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
GACGTGCGCCTGCCAGCCCTGACTGCCCCTTCCCGAGAAGGGTCAAGCTGCCTGGGAAATGCTGCAAGGAG
1810 1820 1830 1840 1850 1860 1870

620 630 640 650 660 670 680
TGGGTGTGTGACGAGCCCAAGGACCAACCGTGGTTGGGCCTGCCCTCGCGGCTTA-CCGAC----TGGGAG
||||| || ||||||| || || ||||||| || || ||||||| || || || || || || || || || || || || ||
TGGGTGTGTGACGAGCCCAAGGACCGCACAGCAGTTGGCCCTGCCCTAGCTGGTGAGTTGTCTCCTTCAAAG
1880 1890 1900 1910 1920 1930 1940

690 700 710 720 730 740 750
ACACGTTTGGCCCAGACCCAACTATGATTAGAGCCAACTGCCTGGTCCAGACCACAGAGTGGAGCGCCTGTT
|| || || || || || || || || || || || || || || || || || || || || || || || || || || || ||
TTAC-TGTTATTCATTCTC-CCCAACTTCAG-GCCAATGCCCAAGTCC--ACCAAATTAAGGGGAAATTGTC
1950 1960 1970 1980 1990 2000 2010

760 770 780 790 800 810 820
CCAAGACCTGTGGGATGGGCATCTCCACCCGGGTT-ACCAATGACAAC--GCCTCCTGCAGGCTAGA--GA
|| || || || || || || || || || || || || || || || || || || || || || || || || || || || ||
CTA--TCCGGATGTTTTACCTTGTGTTTGTGTGTTCTGCTCTCACAGCCTACCGACTGGAAGACACATTTGG
2020 2030 2040 2050 2060 2070 2080

830 840 850 860 870 880 890
AGCAGAGCCGCCTGTGCATGGTCAGGCCTTGCGAAGCTGACCTG-GAAGAGAACATTAAGAAGGGCAAAAAG
||||| || || || || || || || || || || || || || || || || || || || || || || || || || || || || ||
CCCAGACCCAACTATG-ATG--CGAGCC-----AA-CTGCCTGGTCCAGACCACAGAGTGGAGCGCCTGTTC
2090 2100 2110 2120 2130 2140

900 910 920 930 940 950 960
T-GCATCCGTACTCCCAAATCTCCAAGCCTATCAAGTTTGAGCTTTCTGGCTGCA-CCAGCATGAAGACAT
|| || || || || || || || || || || || || || || || || || || || || || || || || || || || ||
TAAGACCTGTGGAATGGGCATCTCCACCCGAGTTACCAATGACAATACCTTCTGCAGACTGGA-GAAG-CAG
2150 2160 2170 2180 2190 2200 2210

970 980 990 1000 1010 1020 1030
A-CCG-AGCTAAATTCTGTGGAGTATGTACCGACGGCCGATGCTGCACCCCCACAGAACCACCACCCTGCC
|| || || || || || || || || || || || || || || || || || || || || || || || || || || || ||
AGCCGCCTCTGCATGGTCAGG-CCCTGCGAAGCTGACC--TGGAGGAAAACATTAAGGTACATCCTCTGCC
2220 2230 2240 2250 2260 2270 2280

1040 1050 1060 1070 1080 1090 1100
GGTGGAGTTCAAGTGCCCTGACGGCGAGGTTCATGAAGAAGAACATGATGT-TCATCAAGACCTGTGCCT---
|| || || || || || || || || || || || || || || || || || || || || || || || || || || || ||
CTAGTCACTCCGTTTTACAGAATGACAGGGAAGAGAACCAGAGCTGGCTGTCTCA-CCTCCCATGTTATTAGA
2290 2300 2310 2320 2330 2340 2350

1110 1120 1130 1140 1150 1160 1170
GCCATTACAACCTGTCCCGGAGACA-ATGAC-ATCT--TTGAATCGCTGTACTACAGGAAGATGTACGGAGAC
|| || || || || || || || || || || || || || || || || || || || || || || || || || || || ||
GGCCTGTTGTCT--CCAGAAATATCTAACCGTGGAGCTGTCTGGCTAGAATGAGAGATGCTGTAACAACAG
2360 2370 2380 2390 2400 2410 2420

1180 1190 1200 1210 1220 1230
ATGGCA---TGAAGCCAGAGAGT---GAGAGACATTAACCTCATTAGACTGGAACTTGAAGTATTACATCT
|| || || || || || || || || || || || || || || || || || || || || || || || || || || || ||
CTGCCAGTTTTTCCACTACAAAATCCCGTGGTGCTAGTTAATTCAA--GACTCCAATGAGGC-CATGG
2430 2440 2450 2460 2470 2480 2490

/number=2
/gene="fisp-12"
/number=3
/gene="fisp-12"
/number=3
/gene="fisp-12"
/number=4
/gene="fisp-12"
/number=4
/gene="fisp-12"
/number=5
/gene="fisp-12"

exon

1675..1926

intron

1927..2058

exon

2059..2270

intron

2271..2641

exon

2642..4006

BASE COUNT 1061 a 1039 c 984 g 1044 t
ORIGIN chromosome 10

Initial Score = 498 Optimized Score = 954 Significance = 0.00
Residue Identity = 50% Matches = 1104 Mismatches = 868
Gaps = 211 Conservative Substitutions = 0

```

      X      10      20      30      40      50
      CCGGGCCGACAGCCCCG-AGACGACAGCCCGGCGC-GTCCCGGTCC-CCACCT-CCG-ACCA
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CAGCCGAAGCAGCGCCGCACTGCCCCGCCGGCGTGAGCCTGGTGCTGGACGGCTGCGGCTGCTGCCGCGTCT
1250 X      1260      1270      1280      1290      1300      1310

      60      70      80      90      100      110      120
      CCGCC-AGC-GCTCCAGGCCCGCGCTCCCCGCTCGCCGCCACCGCGCCCTCCGCTCCGCCCGC-AGTGCCA
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      GCGCCAAGCAGCTGGGAGAACTGTG--TACGGAGCGTGACCCCTGCGACCCACACAAGGGCCTCTTCTGCGA
      1320      1330      1340      1350      1360      1370      1380

      130      140      150      160      170      180      190
      -ACCATGACCGCCGCCAGTATGGGCCCGTCCGCGT-CGCCTTCGT--GGTCCTCCTCGCCCTCTGCAGCCG
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      TTTCGGCTCCCCCGCCA--ACCG-CAAGATTGGAGTGTGCACTGGTAAGACCCTCAGC-CCCATTCAGCCC
      1390      1400      1410      1420      1430      1440      1450

      200      210      220      230      240      250      260
      GCCGGCCGTCGGCCAGAACTGCAGCGGGCCGTGCCGGTGCC-CGGACGAGCCGGCGCCGCGCTGCCCGGCGG
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      CCTTTGCAGAGGCCTCACCTTTTG-GTGTTGGACCACCACCTCTCTCAAGTC--CAGCGTGATACCC-TCTA
      1460      1470      1480      1490      1500      1510

      270      280      290      300      310      320      330
      GCGTGAGCCTCGTGCTGGACGGCTGCGGCTGCTGCCGCGTCTGCGCCAAGCAGCTG-GGCGAGCTGTGCACC
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      GAAAAAGAAAAGCCCCTATCCGCAGCTGCTTCCAACCGGCCCCCTGCAGTC--CTGACCCTAGCTCGTCACC
1520      1530      1540      1550      1560      1570      1580

      340      350      360      370      380      390      400
      GAG-CGCGACCCCTGCGACCCGACACAAGGGCCTCTT--CTGTGA--CTTCGGCTCCCCGGCCAACCGCAAGA
      | | | | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      TTGACATGTACAGTGATA-TAGCTAGCTGTTCTGATCCCTGTGACCCTACGCCTGACC--TCTACAACTTTG
1590      1600      1610      1620      1630      1640      1650

      410      420      430      440      450      460      470
      TCGGCGTGTGC-ACCGCCAAAGATGGTGCTCCCTGCATCTTCGGTGGTACGGTGTACCGCAGCGGAGAGTCC
      || | | | | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      TCTTCCTCTCCTGCAGCCAAAGATGGTGACCCCTGTGTCTTCGGTGGGTGCGGTGTACCGCAGCGGTGAGTCC
1660      1670      1680      1690      1700      1710      1720      1730
```



```

1770      1780      1790      1800      1810      1820      1830
ATGTAAATATGTGGATATATATATATATGTACAGTATCTA-AGTTAATTTAAAGTCATT
1840      1850      1860      1870      1880      1890      1900
TAAAGTTGTTT---GTGCCTTTTTATTTTTGTTTTTAATGCTTTGATATTTCAATGTTAGCCTCAATTTCTG
:      :      :      :      :      :      :      :      :      :      :      :
TGTTTTTTGTTTTAAGTGCTTTTGGGATTTTAAACTGATAGC-CTCAAACCTCCAA--ACACCAT--AGGTAGG
1840      1850      1860      1870      1880      1890      1900

1910      1920      1930      1940      1950      1960      1970
A-AC--ACCATAGGTAGAATGTAAAGCTTGTCTGATCGTTCAAAGCATGAAATGGATACTTATATGGAAATT
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
ACACGAAGCTTATCTGTGATTCAAAACAAAGGAGATACTGCAGTG---GGAATTGTGACCT-GAGTGACTCT
1910      1920      1930      1940      1950      1960      1970

1980      1990      2000      2010      2020      2030
CTGCTCAG---ATAGAATGACAGTCC--GTCAAACAGAT-TGT-TTGCAAAGGGGAGGCATC-AGT-GTCT
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
CTG-TCAGAACAAACAAATGCTGTGCAGGTGATAA-AGCTATGTATTGGAA--GTCAGATTTCTAGTAGGAA
1980      1990      2000      2010      2020      2030      2040

2040      2050      2060      2070      X
TGGCAGGCTGAT--TTCTAGGT-AGGAAATGTGGTAGCTCACG
: : : : : : : : : : : : : : : : : : : : : : : : : :
ATGTGGTCAAATCCCTGTTGGTGAACAAATGGCCTTTATTAAGAAATGGCTGG
2050      2060      2070      2080      X      2090

```

2. US-07-752-427-1 (1-2075)

MUSFISP12A Mouse FISP-12 protein (fisp-12) gene, complete cds

LOCUS	MUSFISP12A	4128 bp ds-DNA	ROD	08-JUL-1991
DEFINITION	Mouse FISP-12 protein (fisp-12) gene, complete cds.			
ACCESSION	M70641			
KEYWORDS	cysteine-rich protein; growth factor-inducible gene.			
SOURCE	Mus musculus (sub_species domesticus) liver/kidney DNA.			
ORGANISM	Mus musculus			
	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;			
	Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.			
REFERENCE	1 (bases 1 to 4128)			
AUTHORS	Ryseck,R.-P., MacDonald-Bravo,H., Mattei,M.-G. and Bravo,R.			
TITLE	Structure, mapping and expression of fisp-12, a growth factor			
	inducible gene encoding a secreted cysteine-rich protein			
JOURNAL	Cell Growth Differ. 2, 225-233 (1991)			
STANDARD	full staff_entry			
FEATURES	Location/Qualifiers			
CDS	join(1052..1114,1202..1424,1675..1926,2059..2270,2642..2938)			
	/product="FISP-12 protein"			
	/gene="fisp-12"			
	/codon_start=1			
mRNA	join(833..1114,1202..1424,1675..1926,2059..2270,2642..4006)			
	/gene="fisp-12"			
TATA_signal	798..804			
	/gene="fisp-12"			
exon	833..1114			
	/number=1			
	/gene="fisp-12"			
intron	1115..1201			
	/number=1			
	/gene="fisp-12"			
exon	1202..1424			
	/number=2			
	/gene="fisp-12"			
intron	1425..1674			

1010 1020 1030 1040 1050 1060 1070

1070 1080 1090 1100 1110 1120 1130 1140

AAGAAGACATGATGTTTCATCAAGACCTGTGCCTGCCATTACAAGTGTCCCGGAGACAATGACATCTTTGAA

1080 1090 1100 1110 1120 1130 1140

AAAAAGAATATGATGTTTCATCAAGACCTGTGCCTGCCATTACAAGTGTCCCGGAGACAATGACATCTTTGAG

1150 1160 1170 1180 1190 1200 1210

TCGCTGTACTACAGGAAGATGTACGGAGACATGGCATGAAGCCAGAGAGTGAGAGACATTAAGTCAATTAGAC

1150 1160 1170 1180 1190 1200 1210

TCCCTGTACTACAGGAAGATGTACGGAGACATGGCGTAAGCCAGGAAGTAAGGGACACGAAGTCAATTAGAC

1220 1230 1240 1250 1260 1270 1280

TGGAAGTGAAGTGAATTCACATCTCATTT--TTCCGTAAAAATGATTTTCAGTAGCACAGTTATTTAAATCT

1220 1230 1240 1250 1260 1270 1280

TATAAGTGAAGTGAATTCACATCTCATTTTCTTCTGTAAAAACAATTACAGTAGCAC-ATTAATTTAAATCT

1290 1300 1310 1320 1330 1340

GTTTTTCTAAGT--GGGGGA-AAAGATTCCCAACCAATTCAAAACATTGTGCCATGTCAAACAAATAGT-C

1290 1300 1310 1320 1330 1340 1350

GTGTTTTTAACTACCGTGGGAGGAACTATCCCAACCAAGTGAGAACGTTATGTCATGGCCA--TACAAGTAG

1350 1360 1370 1380 1390 1400 1410

TATCTTCCCAAGACACTGGTTTGAAGAATGTTAAGAC--TTGACAGTGGG-ACT--ACATTAGTACACAGC

1360 1370 1380 1390 1400 1410 1420

TCTGTCAACCTCAGAC--ACTG--GTTTCGAGACAGTTTACACTTGACAGTTGTTTCAATTAGCGCACAGT

1420 1430 1440 1450 1460 1470 1480

ACCAGAATGTATATTAAGGTGTGGCTTTAGGAGCAGTGGGAGGGTACCGGCCCGGTTAGTATCATCAGATCG

1430 1440 1450 1460 1470 1480

GCCAGAACGCACACTGAGGTGAGTCTCCTGGAACAGT-GGAG-ATGCCAG--GAGAAAGAAAGA-CAGGT--

1490 1500 1510 1520 1530 1540 1550

ACT--CTTATACGAGTAATATGCCTGCTATTTGAAGTGTAAATTGAGAAGGAAAATTTTAGCGTGCTCACTGA

1490 1500 1510 1520 1530 1540 1550

ACTAGCTGAGGTTATTTTAAAGCAGCAGTGTGCCTACTTTTGGAGTGTAAACCGGGGAGGGAAATTA-TAG

1560 1570 1580 1590 1600 1610 1620

CCTGCCTGTAGCC--CCAGTGACAGCTAG-GATGTGCAT-TCTCCAGC-CATCAAGAGACTGAGTCAAGTT

1560 1570 1580 1590 1600 1610 1620 1630

CATGCTTGACAGACAGACCTGCTCTAGCGAGAGCTGAGCATGTGTCCTCCACTAGATGAGGCTGAGTCCAGCT

1630 1640 1650 1660 1670 1680 1690

GTTCTTAAGTCAGAACAGCAGACTCAG-CTCTGA-CATTCTGATTCTGAATGACACTGTTTCAAGGAATCGGAA

1640 1650 1660 1670 1680 1690

GTTCTTTA--AGAACAGCAGTTTCAGCCTCTGACCATTTCTGATTCCAGTGACACTTGTGAGGAGTCAGAG

1700 1710 1720 1730 1740 1750 1760

TCCTGTGCTAGTACTGGACAGCTTGTGGCAAGTGAATTTGCCTGTAAACAAGCCAGATTTTAAATTTAT

1700 1710 1720 1730 1740 1750 1760

CCTTGTCTGTTAGTACTGGACAGCTTGTGGCAAGTGAATTTGCCTGTAAACAAGCCAGATTTT--ATTGAT

1770 1780 1790 1800 1810 1820 1830

ATTGTAAATATTGTGTGTGTGTGTGTGTGTGTATATATATATA--TATGTACAGTT-ATCTAAGTTAATT

210 220 230 240 250 260 270 280

280 290 300 310 320 330 340
GTGCTGGACGGCTGCGGCTGCTGCCGCGTCTGCGCCAAGCAGCTGGGCGAGCTGTGCACCGAGCGCGACCCC
GTGCTGGACGGCTGCGGCTGCTGCCGCGTCTGCGCCAAGCAGCTGGGCGAGCTGTGCACCGAGCGCGACCCC
290 300 310 320 330 340 350

350 360 370 380 390 400 410 420
TGCGACCCGACACAAGGGCCTCTTCTGTGACTTCGGCTCCCCGGCCAACCGCAAGATCGGGCGTGTGCACCGCC
TGCGACCCGACACAAGGGCCTCTTCTGTGACTTCGGCTCCCCGGCCAACCGCAAGATCGGGCGTGTGCACCGCC
360 370 380 390 400 410 420

430 440 450 460 470 480 490
AAAGATGGTGCTCCCTGCATCTTCGGTGGTACGGTGTACCGCAGCGGAGAGTCCTTCCAGAGCAGCTGCAAG
AAAGATGGTGCTCCCTGCATCTTCGGTGGTACGGTGTACCGCAGCGGAGAGTCCTTCCAGAGCAGCTGCAAG
430 440 450 460 470 480 490

500 510 520 530 540 550 560
TACCAGTGACGCTGCCTGGACGGGGCGGTGGGCTGCATGCCCTGTGCAGCATGGACGTTCTGTCTGCCCAGC
TACCAATGCACTTGCCTGGATGGGGCCGTGGGCTGCGTGCCCTATGCAGCATGGACGTTCTGTCTGCCCAGC
500 510 520 530 540 550 560

570 580 590 600 610 620 630
CCTGACTGCCCTTCCCGAGGAGGGTCAAGCTGCCCGGGAAATGCTGCGAGGAGTGGGTGTGTGACGAGCCC
CCTGACTGCCCTTCCCGAGGAGGGTCAAGCTGCCCGGGAAATGCTGCGAGGAGTGGGTGTGTGACGAGCCC
570 580 590 600 610 620 630 640

640 650 660 670 680 690 700
AAGGACCAACCGTGGTTGGGCCTGCCCTCGCGGCTTACCGACTGGAAGACACGTTTGGCCCAGACCCAACT
AAGGACCGCACAGCAGTTGGCCCTGCCCTAGCTGCCTACCGACTGGAAGACACATTTGGCCCAGACCCAACT
650 660 670 680 690 700 710

710 720 730 740 750 760 770 780
ATGATTAGAGCCAACTGCCTGGTCCAGACCACAGAGTGGAGCGCCTGTTCCAAGACCTGTGGGATGGGCATC
ATGATTAGAGCCAACTGCCTGGTCCAGACCACAGAGTGGAGCGCCTGTTCCAAGACCTGTGGGATGGGCATC
720 730 740 750 760 770 780

790 800 810 820 830 840 850
TCCACCCGGGTTACCAATGACAACGCCTCCTGCAGGCTAGAGAAGCAGAGCCGCCTGTGCATGGTCAGGCCT
TCCACCCGAGTTACCAATGACAATACCTTCTGCAGACTGGAGAAGCAGAGCCGCCTGTGCATGGTCAGGCCT
790 800 810 820 830 840 850

860 870 880 890 900 910 920
TGCGAAGCTGACCTGGAAGAGAACATTAAGAAGGGCAAAAAGTGCATCCGTACTCCCAAATCTCCAAGCCT
TGCGAAGCTGACCTGGAAGAGAACATTAAGAAGGGCAAAAAGTGCATCCGTACTCCCAAATCTCCAAGCCT
860 870 880 890 900 910 920

930 940 950 960 970 980 990
ATCAAGTTTGAGCTTTCTGGCTGCACCAGCATGAAGACATACCGAGCTAAATTCTGTGGAGTATGTACCGAC
GTCAAGTTTGAGCTTTCTGGCTGCACCAGCATGAAGACATACCGAGCTAAATTCTGTGGAGTATGTACCGAC
930 940 950 960 970 980 990 1000

1000 1010 1020 1030 1040 1050 1060
GGCCGATGCTGCACCCCCACAGACACCACCCTGCCGGTGGAGTTCAAGTGGCCCTGACGGCGAGGGTCATG
GGCCGATGCTGCACCCCCACAGACACCACCCTGCCGGTGGAGTTCAAGTGGCCCTGACGGCGAGGGTCATG

8.	HSSRICP18	Pseudorabies virus ICP18.5 ge	2524	168	846	7.70	0
9.	HUMCANPRA	Human calcium-dependent prote	1154	164	390	7.46	0
10.	TTHCAAA	T. thermophilus cytochrome caa	1235	164	286	7.46	0
11.	XELRGEE83	x. laevis rrna external transc	865	162	210	7.34	0
12.	XELRGMN3	x. laevis 28s rrna non-transcr	1151	162	243	7.34	0
13.	HS5HCMVCG	Human Cytomegalovirus Strain	229354	158	852	7.10	0
**** 6 standard deviations above mean ****							
14.	HUMRET5	Human mRNA for ret proto-onco	989	154	431	6.86	0
15.	SERERYFGH	S. erythraea 6-deoxyerythromol	2243	152	394	6.74	0
16.	MUSHTF9	Mouse DNA for GC rich region	3725	151	869	6.69	0
17.	HVBLE	Barley (H. vulgare) lectin mRN	972	150	405	6.63	0
18.	BLYLEC	Barley root-specific lectin m	972	150	405	6.63	0
19.	WHTAGGTA	Wheat (T. aestivum) germ agglu	998	149	417	6.57	0
20.	HUMVIM	Human vimentin gene, complete	1749	148	717	6.51	0

Query sequence being compared:US-07-752-427-1 (1-2075)

Number of sequences optimized: 69

Results of the optimized comparison of US-07-752-427-1 (1-2075) with:
Data bank : GenBank 69, all entries
Data bank : UEMBL 28_69, all entries

PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	11
Gap size penalty	0.33		
Cutoff score	124		
Randomization group	0		
Initial scores to save	20	Alignments to save	10
Optimized scores to save	20	Display context	10

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	820	821	0.00
Times:	CPU	Total Elapsed	
	00:00:09.98	00:00:16.00	

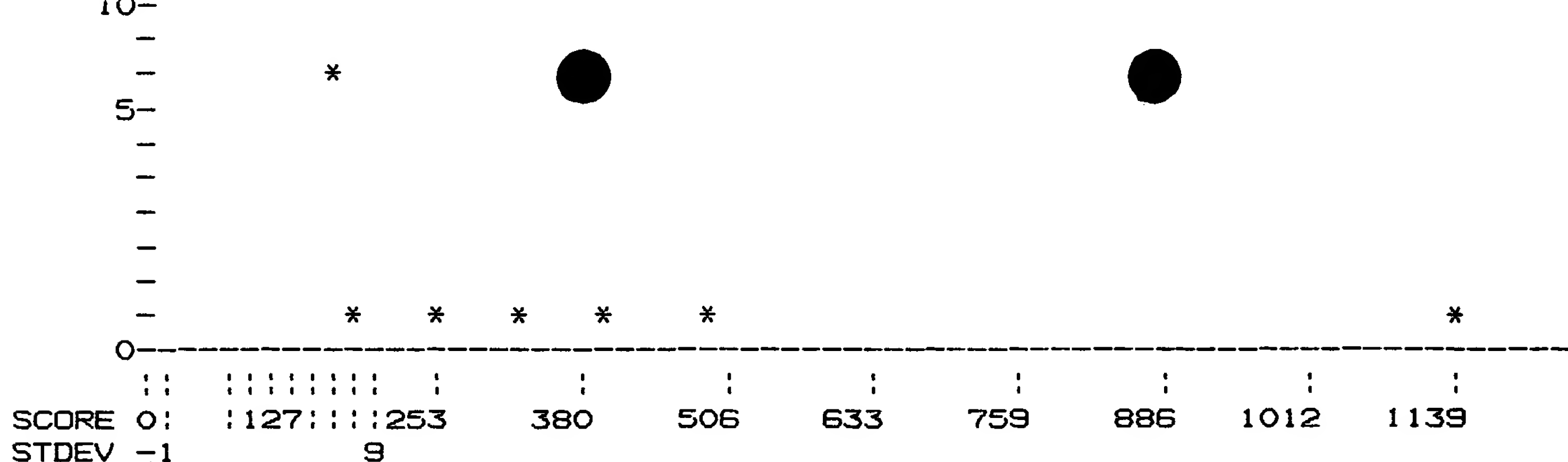
Number of residues: 458541
Number of sequences optimized: 69

The scores below are sorted by optimized score.
Significance is calculated based on optimized score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. MUSFISP12B	Mouse FISP-12 protein (fisp-1	2267	1139	1522	0.00	0
2. MUSFISP12A	Mouse FISP-12 protein (fisp-1	4128	498	954	0.00	0
3. HUMFGFAA	Human fibroblast growth facto	3901	146	878	0.00	0
4. HUMERP	Human erythropoietin gene, co	3398	128	877	0.00	0
5. MUSHTF9	Mouse DNA for GC rich region	3725	151	869	0.00	0
6. HUMSRF	Human serum response factor (420	134	868	0.00	0
7. MUSERPA	Mouse erythropoietin gene, co	3891	139	867	0.00	0
8. CHKCEF	Chicken CEF-10 protein mRNA,	1805	326	865	0.00	0



PARAMETERS

Similarity matrix	Unitary	K-tuple	4
Mismatch penalty	1	Joining penalty	30
Gap penalty	1.00	Window size	11
Gap size penalty	0.33		
Cutoff score	124		
Randomization group	0		
Initial scores to save	20	Alignments to save	10
Optimized scores to save	20	Display context	10

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	39	38	16.75
Times:	CPU	Total Elapsed	
	01:04:13.07	01:15:09.00	
Number of residues:	74447152		
Number of sequences searched:	57272		
Number of scores above cutoff:	69		

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
1. MUSFISP12B	**** 65 standard deviations above mean **** Mouse FISP-12 protein (fisp-1	2267	1139	1522	65.66	0
2. MUSFISP12A	**** 27 standard deviations above mean **** Mouse FISP-12 protein (fisp-1	4128	498	954	27.40	0
3. MUSCYR61A	**** 22 standard deviations above mean **** Mouse Cyr61 mRNA, complete cd	2018	408	848	22.02	0
4. CHKCEF	**** 17 standard deviations above mean **** Chicken CEF-10 protein mRNA,	1805	326	865	17.13	0
5. MMCYR61G	**** 13 standard deviations above mean **** Mouse growth factor inducible	5196	264	833	13.43	0
6. HUMNFIL6	**** 8 standard deviations above mean **** Human gene for nuclear factor	1910	181	621	8.48	0
7. PSEPPI	**** 7 standard deviations above mean **** P.aeruginosa pmi gene encodin	1990	170	718	7.82	0

> O <
O: !O IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Results file us-07-752-427-1.res made by maryh on Thu 6 Feb 92 16:25:14-PST.

Query sequence being compared: US-07-752-427-1 (1-2075)
Number of sequences searched: 57272
Number of scores above cutoff: 69

Results of the initial comparison of US-07-752-427-1 (1-2075) with:
Data bank : GenBank 69, all entries
Data bank : UEMBL 28_69, all entries

100000-
-
N -
U50000-
M -
B -
E -**
R -
-
O -
F10000-
- *
S -
E 5000-
Q -
U *
E -
N - *
C -
E -
S 1000-
-
- *
500-
-
-
- *
100-
-
50-
- *
-
- *
-
-

140	150	160	170	180	190	
LCSMDVRLPSPD---	CPFPRRV+	GKCCEEWVCDEPKD	QTVVGPAL	AAY	-LEDTFGPDPTMIRA---	NC
:	:	:	:	:	:	:
LGMTDYL	VIVEDDDSAI	IPCRTTDP----	ETPV--	TLHNSEGVVPASY	DSRGGFNGTF	TVGPYICEATVKGK
140	150	160	170	180	190	

200	210	220	230	240	250	260
LVQT--	TEWSACSKTCGMG	ISTRVTNDNASCR	LEKQSR	LCMVRPCEAD	LEENIKKGKKC	IRTPKISKPIKFE
:	:	:	:	:	:	:
KFQT	IPFNVYALKAT----	SELDLEMEALK	TVYK-S	GETIVVTC	AVFNNEVVDL	QWTYPGEVKGKGITMLE
200	210	220	230	240	250	260

270	280	290	300	310	320	330
LSGCTSMK---	TYRAKFCGVCTD	GRCCTPHRTT	TLPV-	EFKCPDGEVM	KKNMFIKTC	ACHYNCPGDNDIFE
:	:	:	:	:	:	:
EIKVPSIKL	VYTLTVPEATV	KDSGDYEC	AARQATRE	VKEMKKVT	ISVHEKGF	IEIKPTFSQLEAVNLHEVKH
270	280	290	300	310	320	330

340	X
SLYYRKMYGDMA	
:	
FVVEVRAYPPPRISWLKNNLTL	
340	X 350

> 0 <..

0: 10 IntelliGenetics

> ◻ <

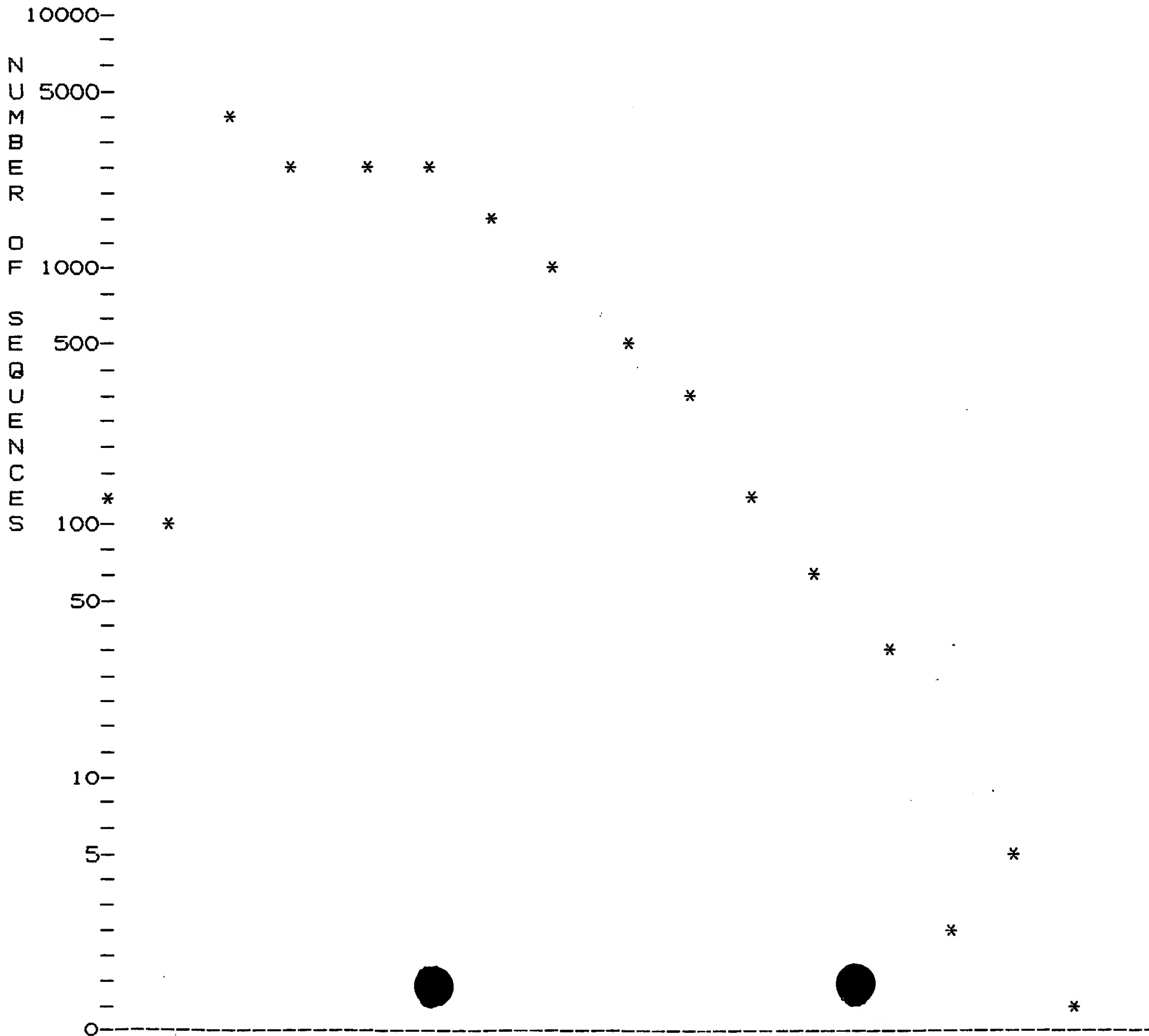
Query sequence being compared: US-07-752-427-2 (1-349)

14140

3377

Results of the initial comparison of US-07-752-427-2 (1-349) with:

Data bank : A-GeneSeq 5, all entries



SCORE 0:	2	3	5	7	8	10	12	13	15
STDEV -1	0			2		3	4	5	

PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	11
Gap size penalty	0.05		
Cutoff score	6		
Randomization group	0		

Initial scores to save	20	Alignments to save	10
Optimized scores to save	20	Display context	10

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	4	5	2.04

Times:	CPU	Total Elapsed
	00:00:42.09	00:01:27.00

Number of residues:	2168208
Number of sequences searched:	14140
Number of scores above cutoff:	3377

The scores below are sorted by initial score.
Significance is calculated based on initial score.

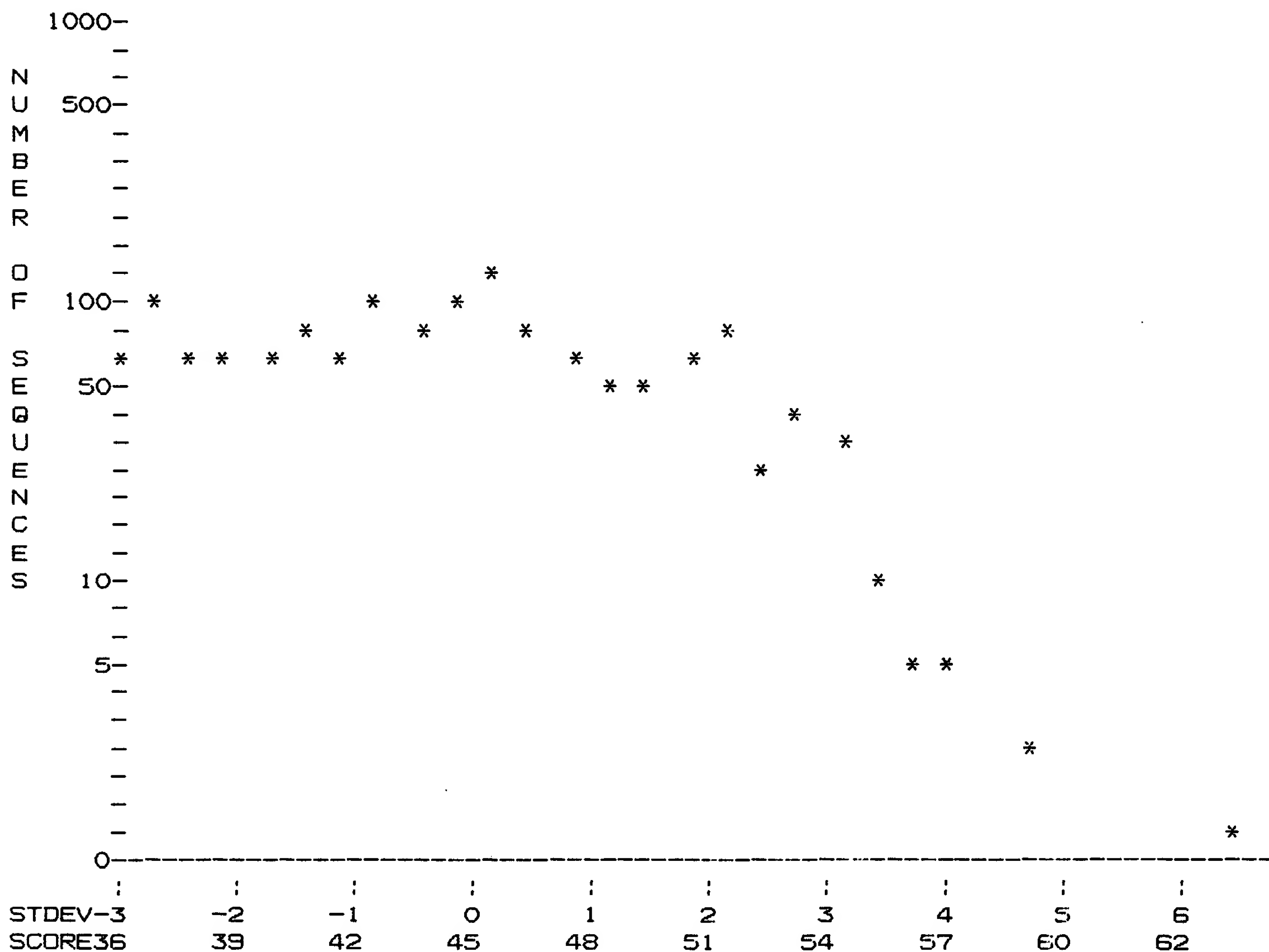
A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 5 standard deviations above mean ****						
1. P93284	Sequence of clone HIV-2 SBL/I	3025	15	57	5.39	0
**** 4 standard deviations above mean ****						
2. P81243	Human spleen trypsin III (try	247	14	37	4.90	0
3. R11741	Granulocyte colony stimulat	783	14	32	4.90	0
4. R11742	Clone 25-1 encoded human G-CS	801	14	35	4.90	0
5. R08039	Cathepsin D fragment of fibro	993	14	45	4.90	0
6. P70373	Human fibronectin gene produc	2327	14	45	4.90	0
7. R12185	Protease inhibitor with varia	114	13	17	4.41	0
8. P94681	Amino acid sequence encoded b	348	13	25	4.41	0
**** 3 standard deviations above mean ****						
9. R10109	Trigramin-beta 1	72	12	20	3.92	0
10. P91320	New trigramin peptide isolate	72	12	20	3.92	0
11. R10110	Trigramin-beta 2.	73	12	20	3.92	0
12. R10106	Trigramin-gamma.	73	12	19	3.92	0
13. P96395	Albolabrin.	73	12	19	3.92	0
14. R06395	Albolabrin.	73	12	19	3.92	0
15. R07329	Amb a I/Antigen E encoded by	92	12	20	3.92	0
16. R07330	Amb a I/Antigen E encoded by	94	12	21	3.92	0
17. R07455	N-terminal sequence of anti-p	120	12	22	3.92	0
18. P91901	Deduced sequence of porcine e	203	12	34	3.92	0
19. P91376	Porcine endothelin III.	203	12	34	3.92	0
20. P90502	Gp.B Eimeria tenella immunoge	237	12	36	3.92	0

Query sequence being compared:US-07-752-427-2 (1-349)

Results of the optimized comparison of US-07-752-427-2 (1-349) with:
Data bank : A-GeneSeq 5, all entries



PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	11
Gap size penalty	0.05		
Cutoff score	6		
Randomization group	0		
Initial scores to save	20	Alignments to save	10
Optimized scores to save	20	Display context	10

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	45	49	3.08
Times:	CPU	Total Elapsed	
	00:01:28.89	00:02:56.00	

Number of residues: 1361708
Number of sequences optimized: 3377

The scores below are sorted by optimized score.

Significance is calculated based on optimized score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Init. Score	Opt. Score	Sig.	Frame
**** 5 standard deviations above mean ****						
1. P92069	Human muscarinic acetylcholin	2135	10	63	5.84	0
**** 4 standard deviations above mean ****						
2. P60053	Sequence of von Willebrand fa	2813	11	59	4.54	0
3. P60462	Sequence of human von Willebr	2813	7	59	4.54	0
**** 3 standard deviations above mean ****						
4. P93284	Sequence of clone HIV-2 SBL/I	3025	15	57	3.89	0
5. P80674	NtrA gene product.	523	6	57	3.89	0
6. R10534	Human 160KD mediator of infla	1427	10	57	3.89	0
7. P93357	Sequence of the catalytic dom	1522	9	57	3.89	0
8. R12108	N-terminal deleted adenyl cyc	1445	9	56	3.57	0
9. R08267	Platelet derived growth facto	1089	6	56	3.57	0
10. R06910	Alpha type PDGF receptor dedu	1089	6	56	3.57	0
11. P94365	Sequence of part of adenylate	1705	9	56	3.57	0
12. R07683	gag protein precursor, p75gag	636	6	56	3.57	0
13. P80810	Sequence of pol protein of HI	1014	6	55	3.25	0
14. P90599	Human retinoblastoma.	970	6	55	3.25	0
15. R07713	Human low density lipoprotein	800	8	55	3.25	0
16. P60057	Factor IX/Factor VII fusion p	453	9	55	3.25	0
17. P90180	Tissue plasminogen activator	534	6	55	3.25	0
18. P60056	Factor VII peptide encoded by	466	9	55	3.25	0
19. R08031	Adenyl cyclase from Bordetell	1706	9	55	3.25	0
20. P60055	Partial Factor VII peptide.	371	9	55	3.25	0

1. US-07-752-427-2 (1-349)

P92069 Human muscarinic acetylcholine m2 receptor gene.

ID P92069 standard; protein; 2135 BP.
AC P92069;
DT 27-SEP-1989 (first entry)
DE Human muscarinic acetylcholine m2 receptor gene.
KW Muscarinic acetylcholine receptor; drug screening; probes; m2; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 120..1517
FT /*tag= a
FT /product=MAR subtype m2
FT misc_signal 74
FT /*tag= b
FT /label=splice acceptor site
FT /note=defines the 5' end of the exon
FT polyA_site 74..2096
FT /*tag= c
PN US7241971-A.
PD 14-MAR-1989.
PF 08-SEP-1988; 241971.
PR 08-SEP-1988; US-241971.
PA (USSH) Nat Inst of Health.
DR WPI; 89-165452/22.
DR P-PSDB; P96203.
PT Cloned genes for muscarinic acetylcholine receptors -
PT for drug screening and diagnostic use.
PS Disclosure; p; English.
CC The sequence may be useful for synthesis of hybridisation probes for
CC diagnostic use. The genes are cloned by screening a rat cerebral cortex

CC cDNA library with a probe based on nucleotides 170-223 of the non-coding
 CC strand of porcine brain (ml) cDNA (modified at positions 5, 38 and 53);
 CC identifying cDNA clones by hybridisation with BamHI or XhoI digests of
 CC plasmid DNA from 12 cultures contg. up to 500000 independent clones;
 CC reculturing until a pool contg. less than 5000 indep. clones with a
 CC single hybridisation band are identified, and isolating individual clones
 CC by colony hybridisation.
 CC Stable cell lines are produced by transfecting Chinese hamster ovary cells
 CC (CHO-K1) with various pCD vectors contg. the gene inserts.
 CC See also N92068-N92067 and N90086.
 SQ Sequence 2135 BP; 629 A; 462 C; 448 G; 596 T;

Initial Score = 10 Optimized Score = 63 Significance = 5.84
 Residue Identity = 20% Matches = 75 Mismatches = 266
 Gaps = 20 Conservative Substitutions = 0

```

      X      10      20      30      40      50      60
      MTAASMGPVRVAFVLLALCSRPAVGQNCSGPCRCPPDEPAPRCPAGVSLVLDGCGCCRVCAK
      |      |      |      |      |      |      |
AAAGTCAACCGCCACCTCCAGACCGTCAACAATTACTTTTTATTTCAGCTTGGCCTG-TGCTGACCTTATCAT
 270   X   280      290      300      310      320      330

      70      80      90      100      110      120      130
      QLGELCTERDPCDPHKGLFCDFGSPANRKIGVCT--AKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDG--A
      |      |      |      |      |      |      |      |      |      |
AGGTGTTTTCTC-CATGAACTTG-TACACCCTCTACACTGTGATTGGTTACTGGCCTTTGGGACCTGTGGTG
 340      350      360      370      380      390      400

      140      150      160      170      180      190      200
      VGCMP LCSMDVRLPSPDCPFPRRVKLP GKCC EEWVCDEPKDQTVVG--PALAAYRLEDTFGPDPTMIRANCL
      |      |      |      |      |      |      |      |      |      |
TGTGACCTTTGGCTAGCCCTGGACTATGTGGTCAGCAATGCCCTCAGTTATGAATCTGCTCATCAT--CAGC-
 410      420      430      440      450      460      470

      210      220      230      240      250      260      270
      VQTEWSACSKTCGMGISTRVTNDNASCRLEKQSR LCMVRPCEADLEENIKKG--KKCIRTPKISKPIKFEL
      |      |      |      |      |      |      |      |      |      |
-TTTGACAGGTACTTCTGT-GTCACAAAACCTCTGACCTACCCAGTCAAGCGGACCACAAAAATGGCAGGTA
 480      490      500      510      520      530      540

      280      290      300      310      320      330
      SGCTSMKTYRAKFCGVCT-DGRCCTPHRTTTL PVEFKCPDGEVMKKNMMFIKTC---ACHYNCPGDNDIFES
      |      |      |      |      |      |      |      |      |      |
TGATGATTGCAGCTGCCTGGGTCCTCTCTTTCATCCTCTGGGCTCCAGCCATTCTCTTCTGGCAGTTCATTG
 550      560      570      580      590      600      610

340      X
  LYYRKM YGDMA
      |      |
TAGGGGTGAGAACTGTGGAGG
 620      X 630

```

2. US-07-752-427-2 (1-349)

P60053 Sequence of von Willebrand factor (vWF).

ID P60053 standard; Protein; 2813 AA.
 AC P60053;
 DT 22-JUL-1991 (first entry)
 DE Sequence of von Willebrand factor (vWF).
 KW Vascular injury; platelet plug formation.
 OS Homo sapiens.
 PN EP-197592-A.
 PD 15-OCT-1986.
 PF 26-MAR-1986; 200518.
 PR 01-APR-1985; NL-000961.

PA (VR1E-) STICHT VRIEND LANDS.
 PI Pannekoek H; Verwey C; Diergaarde PJ; Hart MHL;
 DR WPI; 86-273504/42.
 DR N-PSDB; N60061.
 PT Recombinant cDNA plasmid or phage - contg. C-DNA fragment which
 PT codes for biological activity of human von Willebrand factor
 PS Disclosure; Fig 3; 37pp; English.
 CC vWF (glyco) protein having the AA sequence corresponding to the
 CC nucleotide sequence of 2518-8667 or 295-2517 of N60061 is claimed.
 CC Also claimed are new microorganisms, animal cell or human cell contg.
 CC the recombinant cDNA plasmid or phage; e.g. strain E.coli DH1 contg.
 CC the recombinant cDNA plasmid pSP8800vWF is deposited as C.B.S. No
 CC 163.86.
 SQ Sequence 2813 AA;
 SQ 160A; 137R; 101N; 160D; 0 B; 217C; 142Q; 181E; 0 Z; 207G; 77 H;
 SQ 97 I; 233L; 104K; 55 M; 93 F; 173P; 201S; 146T; 27 W; 81 Y; 221V;

Initial Score = 11 Optimized Score = 59 Significance = 4.54
 Residue Identity = 20% Matches = 77 Mismatches = 248
 Gaps = 43 Conservative Substitutions = 0

```

      X      10      20      30      40      50
      MTAASMG PVRV-AFV VLLALCS RPAVGQ-NC SGPC-RCPDEPAPRCPAGVSLVLDGCGCCRV
      :      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
530 X 540      550      560      570      580      590
60      70      80      90      100     110     120     130
CAKQLGELCTER-DPCDPHKGLFCDFGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGA
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
AVSPLPYLRNCRYDVCSCSDGRECLCGALA-----SYAAACAGRGVRVAWREPGRCELNCPKGQ-VYLQCG
600      610      620      630      640      650      660
      140      150      160      170      180      190
VGC MPLCSMDVRLPSD-----CPFPRRVKLP GKCC EEWVCDEPKDQTVVGPALAAAYRLEDTFGPDP
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TPCNLTC---RSLSY PDEECNEACLEGCFCP-----PGLYMDERGD CVPKAQCP CYDGEIFQPEDIFSDHH
      670      680      690      700      710      720
      200      210      220      230      240      250      260
TMIRANCLVQTT EWSACSKTCGMGISTRVTNDNASCRLEKQS---RLCMV-RPCEADLEENIKKGKKCIRTP
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
TM---CYCEDGFMHCTMSGVPGSLLPDAVLSSPLSHR-SKRSLSCRPPMVKLVCPAD--NLRAEGLECTKT-
      730      740      750      760      770      780      790
      270      280      290      300      310      320      330
KISKPIKFELSGCTSMKTYRAKFCGVCTDGRCC TPHRTTTL PVEFKCPDGEVMKKNMFIKTCACHYNC-PG
: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CQNYDLECM SMGCVS--GCLCPPGMVRHENRCVALERCPCFHQ GKEYAPGETVKIGCNTCVCRDRKWNCTDH
      800      810      820      830      840      850      860
      340      X
DNDIFESLYYRKMYGDMA
: : : : :
VCDATCSTIGMAHYLTFDGLKYLFPGEC
      870      880
  
```

3. US-07-752-427-2 (1-349)

P60462 Sequence of human von Willebrand Factor (VWF) prec

ID P60462 standard; Protein; 2813 AA.
 AC P60462;
 DT 25-JUN-1991 (first entry)
 DE Sequence of human von Willebrand Factor (VWF) precursor.

KW Chronic renal failure; therapy; factor VIII C.
 OS Homo sapiens.
 PN W08606096-A.
 PD 23-OCT-1986.
 PF 10-APR-1986; U00760.
 PR 11-APR-1985; US-722108.
 PA (CHIL-) CHILDRENS MED CENT.
 PA (GINS/) GINSBURG D.
 PI Ginsburg D, Orkin SH, Kaufman RJ;
 DR WPI; 86-291663/44.
 DR N-PSDB; N60404.
 PT Pure Von Willebrand Factor - produced using an expression vector
 PT including a DNA sequence encoding functional VWF protein
 PS Disclosure; Table 2, Pages 18-36A; 54pp; English.
 CC cDNA clones pVWH33, pVWH5 and PVWE6 which span 9 kb pairs of DNA and
 CC encompass the entire protein coding region of VWF, were selected to
 CC construct full length cDNA (N60404). The pure VWF produced is useful
 CC in the treatment of von Willebrand's disease (VWD) and the patients
 CC with chronic renal failure whose abnormal bleeding times are
 CC corrected by crude cryoprecipitate. Pure VWF can also be used to
 CC carry, stabilise and improve the therapeutic efficacy of factor
 CC VIII:C.
 SQ Sequence 2813 AA;
 SQ 154A; 143R; 98 N; 155D; 0 B; 234C; 133Q; 181E; 0 Z; 205G; 70 H;
 SQ 95 I; 227L; 108K; 56 M; 89 F; 176P; 207S; 151T; 26 W; 79 Y; 226V;

Initial Score	=	7	Optimized Score	=	59	Significance	=	4.54
Residue Identity	=	20%	Matches	=	76	Mismatches	=	259
Gaps	=	28	Conservative Substitutions	=			=	0

	X		10		20		30		40		50		60
	MTAASMG-PVRVAFVLLALCSRPAV-GQNCSGPCRCPPDEPAPRCPPAGVSLVLDGCGCCRV												
	:	:	:	:	:	:	:	:	:	:	:	:	:
CLPDKVCVHRSTIYPVGQFWEEGCDVCTCTDMEDAVMGLRVAQCSQKPCE--DSCRSGFTYVLHEGECCGRC													
2430	X	2440	2450	2460	2470	2480	2490						

	70		80		90		100		110		120	
	AKQLGELCT---ERDPCDPHKGLFCDFGSPAN-RKIGVCTAKDGAPCIFGGTV-YRSGESFQSSCKYQCTCL											
	:	:	:	:	:	:	:	:	:	:	:	:
LPSACEVVTGSPRGDSQSSWKSVGSQWASPENPCLINECVRVKEEVFIQQRNVSCPQLEVPVCPSPGFQLSCK												
2500	2510	2520	2530	2540	2550	2560						

	130		140		150		160		170		180		190
	DGAVGCMPLCSMDVRLPSPDCPFPRRVKLP GKCC EEWVCDEPKDQTVVGPALAAAYRLE---DTFGPDPTMIR												
	:	:	:	:	:	:	:	:	:	:	:	:	:
TSA--CCPSC---RCERMEACMLNGTVIGPGKTVMIDVCTTCRCMVQVG-VISGFKLECRKTTCPNCP--LG													
2570	2580	2590	2600	2610	2620	2630							

	200		210		220		230		240		250		260
	ANCLVQTTTEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKISKPIKF												
	:	:	:	:	:	:	:	:	:	:	:	:	:
YKEENNTGECCGRCLPTACTIQLRGGQIMTLKRDETLQDGCDFHFC-KVNERGEYFWEKRVTGCPPFDEHKC													
2640	2650	2660	2670	2680	2690	2700							

	270		280		290		300		310		320		330
	ELSGCTSMKTYRAKFCGVCTDGRCCCTPHRTTTL-PVEFKCPDGEVMKKNMFIKTCACHYNCPGD-ND--IF												
	:	:	:	:	:	:	:	:	:	:	:	:	:
LAEGGKIMK-IPGTCCDTCEEPEC--NDITARLQYVKVGSKSEVEVDIHYCQGKASKAMYSIDINDVQDQ													
2710	2720	2730	2740	2750	2760	2770							

	340		X
	ESLYYRKMYGDMA		
	:	:	:
CSCCSPTRTEPMQVALHCTNGSV			
2780	X	2790	

4. US-07-752-427-2 (1-349)
P93284 Sequence of clone HIV-2 SBL/ISY.

ID P93284 standard; protein; 3025 AA.
AC P93284;
DT 06-APR-1990 (first entry)
DE Sequence of clone HIV-2 SBL/ISY.
KW HIV-2; proviral clone HIV-2 SBL/ISY;
OS Human immunodeficiency virus 2.
PN US7331212-A.
PD 29-AUG-1989.
PF 31-MAR-1989; 331212.
PR 31-MAR-1989; US-331212.
PA (USSH) US Dept. Health and Human Services.
PI Franchini G, Wong-Staal F, Gallo R;
DR WPI; 89-339698/46.
DR N-PSDB; N92119.
PT Complete human immunodeficiency type 2 proviral clone - used to generate
PT animal model for function studies of HIV genes in vivo.
PS Disclosure; Fig. 5; 43pp; English.
CC The protein is encoded by the second reading frame of HIV-2 SBL/ISY, a
CC proviral clone of HIV-2.
SQ Sequence 3025 AA;
SQ 143 A; 278 R; 126 N; 104 D; 0 B; 96 C; 191 Q; 115 E; 0 Z; 244 G;
SQ 95 H; 153 I; 224 L; 246 K; 65 M; 63 F; 160 P; 291 S; 193 T; 53 W;
SQ 80 Y; 105 V;

Initial Score = 15 Optimized Score = 57 Significance = 3.89
Residue Identity = 20% Matches = 75 Mismatches = 256
Gaps = 41 Conservative Substitutions = 0

```

      X      10      20      30      40      50
      MTAASMGPVRFVLLALCSRPAVGQNCSGPCRCPEPA-----PRCPAGVSLVLDGCGC
      :      :      :      :      :      :      :      :      :
MDSHPETSGCPKLGGTNLPRNKDQTLMA--NRKDDTHGRSTVDRISSRAGGKQNYLKPGTGGTLLPRGKRV
980      990      1000      1010      1020      1030      1040

      60      70      80      90      100      110      120
CRVCAK---QLGELCTERDPCDPHKGLFCDFGSPANRKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCT
:      :      :      :      :      :      :      :      :
RSNSSKGSRQSVDISTPGRKNSKSRKICKDKKYPYQRGQVVG---GSSENKRKRSTSHLGTNTKISPTSRKR
1050      1060      1070      1080      1090      1100      1110

      130      140      150      160      170      180      190
CLDGAVGCMPLCS-MDVRLPSPDCPFRRVKLPKCCCEEWVCDEPKDQTVVGPALAA-----YRLEDTFGPD
:      :      :      :      :      :      :      :      :
DLGTVVG---LLASDMDPRLGLRIHPTVGQVSIPGKRSYTRRRD-----LLHGWILQAIKRGKSRIYNRRKRQ
1120      1130      1140      1150      1160      1170      1180

      200      210      220      230      240      250
PTMIRANCLVQTTTEWSACSKTCGMGISTRVTNDNASCL-----EKQSRLCMVRPCEADLEENIKKG-KKCIR
:      :      :      :      :      :      :      :      :
SKDIRAN--YQPT--SRIRSLCNGSNRLRSKSQYCS-RLTVCNGNSNRPTGIRENSKNYRRDDKKGSNLCCM
1190      1200      1210      1220      1230      1240

      260      270      280      290      300      310      320
TP-KISKPIKFELSGCTS-MKTYRAKFCGVCTDGR-CCTPHRTTTLPEVEFKCPDGEVMKKNMFIKTCACHY
:      :      :      :      :      :      :      :      :
GPGPQRHRRKSRNPLSKSGHQTSIIPRENRRASGRTWKISQCKRTSPVWITQPGGKTNSKHMCPVPTKRGSY
1250      1260      1270      1280      1290      1300      1310

      330      340      X
NCPGDNDIFESLYRKMYGDMA
:      :
-TWASKCRTRHLANGLHTLRMKNHYSSSTCK
```



```

      X      10      20      30      40      50      60
      MTAASMGFVRVAFVLLALCSRPAVGQNCSGPCRCPCDEPAPRCPAGVSLVLDGCGCCRVCAK
      : : : : : : : : : : : : : : : : : :
KRTENVQYRHVELARVGQV-VEVDITLEHVQHIIGGAGNDSITGNAHD----NFLAGGSGDDRDLGGAGNDTL
      730 X      740      750      760      770      780

      70      80      90      100      110      120      130
      QLGELCTERDPCDPHKGLFCDFGSPAN---RKIGVCTAKDGAPCIFGGTVYRSGESFQSSCKYQCTCLDGAV
      : : : : : : : : : : : : : : : : : :
VGGEGQNTVIGGAGDDVFLQDLGVWSNQLDGGAGVDTVKNVHQPSEERLERMG-----DTGIHADLQKGTV
790      800      810      820      830      840      850

      140      150      160      170      180      190
      GCMP---LCSMDVRLPSPDCPFPR---RVKLPGKCCEEWVCDEPKDQTVVG--PALAAYRLEDT-FGPDPTM
      : : : : : : : : : : : : : : : : : :
EKWPALNLFSDHVKNIEENLHGSRLNDRIAGDDQDNELWGHGNDTIRGRGGDDILRGGLGLDTLYGEDGND
      860      870      880      890      900      910      920

      200      210      220      230      240      250      260
      IRANCLVQTTWEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKISKPI
      : : : : : : : : : : : : : : : : : :
I----FLQDDETVSDDIDGGAGLDT-VDYSAMIHPGRIVAPHEYGFGEADLSREWVRKASALGVDYYDNVR
      930      940      950      960      970      980      990

      270      280      290      300      310      320      330
      KFELSGCTSMK---TYRAKFCGVCTDGRCTPHRTTTLPVEFKCPDGEVMKKNMFIKTCACHYNCPGDNDI
      : : : : : : : : : : : : : : : : : :
NVENVIGTSMKDVLIQDAGANTLMGQGGDDT-VRGGDGDDLLFGGDGNDMLYGDAGNDT---LYGGLGD-DT
      1000      1010      1020      1030      1040      1050      1060

      340      X
      FESLYYRKMYGDMA
      : :
      LEGGAGNDWFGQTQAREHDVLRGG
      1070      X 1080
```

9. US-07-752-427-2 (1-349)

R08267 Platelet derived growth factor (PDGF) receptor pro

ID R08267 standard; protein; 1089 AA.
AC R08267;
DT 07-MAR-1991 (first entry)
DE Platelet derived growth factor (PDGF) receptor protein.
KW Atherosclerosis; fibrotic diseases.
OS Homo sapiens.
PN W09014425-A.
PD 29-NOV-1990.
PF 21-MAY-1990; U02849.
PR 22-MAY-1989; US-355018.
PA (ZYMO-) ZYMOGENETICS INC.
PI Kelly JD, Murray MJ;
DR WPI; 90-375992/50.
DR N-PSDB; Q06869.
PT DNA encoding platelet-derived growth factor - used to transform
PT cells for culturing to detect PDG agonists and antagonists
PS Claim 1; Fig 1; 30pp; English.
CC Gene product may be expressed from a transformed cell. It has
CC utility in dection of PDGF agonist and antagonist analogues, binding
CC AA, AB and BB isoforms. PDGF agonists may be used to enhance wound
CC healing, and antagonists may be used to block the effects of PDGF
CC eg. in treatment of atherosclerosis or fibrotic diseases.
SQ Sequence 1089 AA;
SQ 54 A; 44 R; 44 N; 66 D; 0 B; 17 C; 25 Q; 94 E; 0 Z; 58 G; 25 H;

SW 71 1; 105L; 69 K; 25 M; 32 F; 51 P; 93 S; 68 T; 13 W; 47 Y; 88 V;

Initial Score = 6 Optimized Score = 56 Significance = 3.57
Residue Identity = 19% Matches = 72 Mismatches = 251
Gaps = 44 Conservative Substitutions = 0

```

      X   10           20           30           40           50           60
MTAASMGPVRFVAFVVL-LALCSRPAVGQNCSGPCRCPCDEPAFRCPCAGVSLVLDGCGCCRV-----CAKQLGEL
  ::      :: ::      :      :      :      :      :      :      :      :
      X   10           20           30           40           50           60
MGTSHPAFLVLGCLLTGLSLILCQLSLPSILPNENKVVQLNSSFSLRFCGESEVSWQYPMSEEESS

      70           80           90          100          110          120          130
CTERDPCDPHKGLFCDFGSPANRKIGVCTAKDGAPCIFGGTVY-----RSGESFQSSCKYQCTCLDGAVGCMF
  :      : ::      :      :      :      :      :      :      :      :
DVEIRNEENNNSGLF-----VTLLEVSSA--SAAHTGLYTCYYNHTQTEENELEGRHIYIYVPDPDVAFP
      70           80           90          100          110          120          130

      140          150          160          170          180          190
LCSMDVRLPSPD---CPFFRRVKLPQKCCCEEWVCDEPKDQTVVGPALAAAYR--LEDTFGPDPTMIRA---NC
  :      :      :      :      :      :      :      :      :      :
LGMTDYLIVIVEDDDSAIIPCRTTDP---ETPV--TLHNSEGVVPASYDSRQGFNGTFTVGPYICEATVKGK
      140          150          160          170          180          190

200           210          220          230          240          250          260
LVQT--TEWSACSKTCGMGISTRVTNDNASCRLEKQSRLCMVRPCEADLEENIKKGKKCIRTPKISKPIKFE
  ::      :      :      :      :      :      :      :      :      :      :
KFQTIPFNVYALKAT-----SELDLEMEALKTVYK-SGETIVVTCAVFNNEVVDLQWTYPGEVKKGKITILE
      200          210          220          230          240          250          260

270           280          290          300          310          320          330
LSGCTSMK---TYRAKFCGVCTDGRCTPHRTTTLTPV-EFKCPDGEVMKKNMFIKTCACHYNCPGDNDIFE
  :      :      :      :      :      :      :      :      :      :
EIKVPSIKLVYTLTVPEATVKDSGDYECARQATREVKEMKKVTISVHEKGFIEIKPTFSQLEAVNLHEVKH
      270          280          290          300          310          320          330

340           X
SLYYRKMYGDMA
      :
FVVEVRAYPPPRISWLKNNLTL
      340           X 350
```

10. US-07-752-427-2 (1-349)

R06910 Alpha type PDGF receptor deduced from TR4 cDNA clo

ID R06910 standard; protein; 1089 AA.
AC R06910;
DT 16-JAN-1991 (first entry)
DE Alpha type PDGF receptor deduced from TR4 cDNA clone.
KW Platelet derived growth factor; T11.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 1..23
FT /label=signal peptide
FT Domain 24..524
FT /label=ligand binding domain
FT Domain 525..548
FT /label=transmembrane region
FT Domain 549..599
FT /label=juxtamembrane domain
FT Binding-site 600..627
FT /label=ATP binding site
FT Modified-site 849
FT /label=tyrosine autophosphorylation site
FT Modified-site 42..44

FT /label=N-glycos_site
 FT Modified-site 76..78
 FT /label=N-glycos_site
 FT Modified-site 103..105
 FT /label=N-glycos_site
 FT Modified-site 179..181
 FT /label=N-glycos_site
 FT Modified-site 353..355
 FT /label=N-glycos_site
 FT Modified-site 359..361
 FT /label=N-glycos_site
 FT Modified-site 458..460
 FT /label=N-glycos_site
 FT Modified-site 468..470
 FT /label=N-glycos_site
 PN W09010013-A.
 PD 07-SEP-1990.
 PF 08-FEB-1990; U00617.
 PR 09-FEB-1989; US-308282.
 PA (USDC) US SEC OF COMMERCE.
 PI Matsui T, Aaronson SA, Pierce JH;
 DR WPI; 90-290306/38.
 DR N-PSDB; 005989.
 PT Type alpha platelet-derived growth factor receptor gene - useful
 PT for transforming cells to express novel protein receptor and also
 PT susceptible to genetic engineering.
 PS Claim 7; Fig 3; 64pp; English.
 CC The TR4 clone is the largest cDNA clone related to the T11 genomic
 CC clone, isolated from a library prepd. from human thymus DNA. The
 CC T4 cDNA clone was isolated from a M426 human embryo fibroblast
 CC cDNA library. The coding region can be introduced into the pSV2
 CC gpt vector with a simian sarcoma virus LTR as a promoter and
 CC expressed in a host. The resulting protein is a novel PDGF
 CC receptor designated type alpha (the known receptor is designated
 CC type beta). The polypeptide has a calculated molecular mass of 120
 CC kD and has all the characteristics of a membrane spanning tyrosine
 CC kinase receptor. The extracellular region comprises a hydrophobic
 CC signal peptide and a ligand binding domain which has structural
 CC homology with the PDGF-R/CSF1-R subfamily. Ten Cys residues are
 CC spaced at the same positions as in other receptors of the sub-
 CC family and eight potential N-linked glycosylation sites are also
 CC present. A hydrophobic segment spans the membrane and the cyto-
 CC plasmic region comprises a juxtamembrane region, a tyrosine kinase
 CC region split into TK1 and TK2 by a hydrophilic interkinase region
 CC and a hydrophilic C-terminal tail. The TK region includes the
 CC consensus ATP binding sequence (G-X-G-X-X-G...K) and a tyrosine
 CC autophosphorylation site homologous to that of pp60(v-src).
 SQ Sequence 1089 AA;
 SQ 54 A; 44 R; 44 N; 66 D; 0 B; 17 C; 26 Q; 93 E; 0 Z; 58 G; 25 H;
 SQ 70 I; 104L; 69 K; 26 M; 32 F; 51 P; 93 S; 68 T; 13 W; 47 Y; 89 V;

Initial Score	=	6	Optimized Score	=	56	Significance	=	3.57
Residue Identity	=	19%	Matches	=	72	Mismatches	=	251
Gaps	=	44	Conservative Substitutions	=			=	0

X	10	20	30	40	50	60
MTAASMG PVRVAFVVL-LALCSRPAVGQNC SGPCRC PDEPAPRC PAQVSLVLDGCGCCRV-----CAKQLGEL						
:	:	:	:	:	:	:
MGTSHPAFLVLGCLLTGLSLILCQLSLPSILPNENEKVVQLNSSFSLR CFGESEVSWQYPMSEEESS						
X	10	20	30	40	50	60

70	80	90	100	110	120	130
CTERDPCDPHKGLFCDFGSPANRKIGVCTAKDGAPCIFGGTVY-----RSGESFQSSCKYQCTCLDGAVGCMF						
:	:	:	:	:	:	:
DVEIRNEENNSGLF-----VTGLEVSSA--SAAHTGLYTCYYNHTQTELENELEGRHIYIYVDPDPDAFVP						
70	80	90	100	110	120	130